



Tsang, C., Boulton, C., Burgon, V., Johansen, A., Wakeman, R., & Cromwell, D. A. (2017). Predicting 30-day mortality after hip fracture surgery: Evaluation of the National Hip Fracture Database case-mix adjustment model. *Bone and Joint Research*, 6(9), 550-556.  
<https://doi.org/10.1302/2046-3758.69.BJR-2017-0020.R1>

Publisher's PDF, also known as Version of record

License (if available):  
CC BY

Link to published version (if available):  
[10.1302/2046-3758.69.BJR-2017-0020.R1](https://doi.org/10.1302/2046-3758.69.BJR-2017-0020.R1)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the final published version of the article (version of record). It first appeared online via Bone and Joint Publishing at <https://online.boneandjoint.org.uk/doi/full/10.1302/2046-3758.69.BJR-2017-0020.R1>. Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>



## ■ HIP

# Predicting 30-day mortality after hip fracture surgery

## EVALUATION OF THE NATIONAL HIP FRACTURE DATABASE CASE-MIX ADJUSTMENT MODEL

**C. Tsang,  
C. Boulton,  
V. Burgon,  
A. Johansen,  
R. Wakeman,  
D. A. Cromwell**

*Clinical Effectiveness Unit, The Royal College of Surgeons of England, London, United Kingdom*

■ C. Tsang, PhD, Honorary Assistant Professor, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK and Honorary Lecturer, The Royal College of Surgeons of England, 35-43 Lincoln's Inn Fields, London WC2A 3PE, UK.

■ C. Boulton, BA, Falls and Fragility Fracture Audit Programme Manager,

■ V. Burgon, BA, National Hip Fracture Database Project Manager,

■ R. Wakeman, FRCS, Clinical Lead, Clinical Effectiveness and Evaluation Unit, Royal College of Physicians, 11 St Andrews Place, London NW1 4LE, UK.

■ A. Johansen, FRCP, Consultant Orthogeriatrician and Clinical Lead for Orthogeriatrics, Trauma Unit, University Hospital of Wales, Heath Park, Cardiff CF14 4XW, UK.

■ D. A. Cromwell, PhD, Professor of Health Services Research, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK and Director of Clinical Effectiveness Unit, The Royal College of Surgeons of England, 35-43 Lincoln's Inn Fields, London WC2A 3PE, UK.

Correspondence should be sent to C. Tsang;  
email: [ctsang@rcseng.ac.uk](mailto:ctsang@rcseng.ac.uk)

doi: 10.1302/2046-3758.69.BJR-2017-0020.R1

*Bone Joint Res* 2017;6:550–556.

Received: 28 January 2017

Accepted: 28 May 2017

### Objectives

The National Hip Fracture Database (NHFD) publishes hospital-level risk-adjusted mortality rates following hip fracture surgery in England, Wales and Northern Ireland. The performance of the risk model used by the NHFD was compared with the widely-used Nottingham Hip Fracture Score.

### Methods

Data from 94 hospitals on patients aged 60 to 110 who had hip fracture surgery between May 2013 and July 2013 were analysed. Data were linked to the Office for National Statistics (ONS) death register to calculate the 30-day mortality rate. Risk of death was predicted for each patient using the NHFD and Nottingham models in a development dataset using logistic regression to define the models' coefficients. This was followed by testing the performance of these refined models in a second validation dataset.

### Results

The 30-day mortality rate was 5.36% in the validation dataset ( $n = 3861$ ), slightly lower than the 6.40% in the development dataset ( $n = 4044$ ). The NHFD and Nottingham models showed a slightly lower discrimination in the validation dataset compared with the development dataset, but both still displayed moderate discriminative power (c-statistic for NHFD = 0.71, 95% confidence interval (CI) 0.67 to 0.74; Nottingham model = 0.70, 95% CI 0.68 to 0.75). Both models defined similar ranges of predicted mortality risk (1% to 18%) in assessment of calibration.

### Conclusions

Both models have limitations in predicting mortality for individual patients after hip fracture surgery, but the NHFD risk adjustment model performed as well as the widely-used Nottingham prognostic tool and is therefore a reasonable alternative for risk adjustment in the United Kingdom hip fracture population.

**Cite this article:** *Bone Joint Res* 2017;6:550–556

**Keywords:** Hip fracture, Mortality, Orthopaedics, Risk factors, Surgery

### Article focus

- Numerous models, both generic and hip fracture-specific, have been developed to predict mortality after hip fracture surgery.
- The Nottingham Hip Fracture Score has been validated and applied widely in the United Kingdom for predicting mortality in hip fracture patients.
- The refined risk adjustment model used by the National Hip Fracture Database was compared with the Nottingham Hip Fracture Score using national clinical audit data.

### Key messages

- The model used by the National Hip Fracture Database performed as well as the Nottingham Hip Fracture Score.
- The National Hip Fracture Database model is a suitable alternative to the Nottingham Hip Fracture Score for risk adjustment in the United Kingdom hip fracture population.

### Strengths and limitations

- Patients who had any type of surgery for hip fracture were included in the study,

and thus the findings are applicable to the entire hip fracture population.

- Data quality in this study was better than in earlier validation studies of the Nottingham model, and missing data were managed using the robust approach of multiple imputation.
- There may be residual confounding due to risk factors that were not, or were only partly, captured in the dataset.

## Introduction

Over 65 000 people aged 60 years or older suffer a hip fracture in England, Wales and Northern Ireland every year.<sup>1</sup> The injury is associated with increased risk of death, with only approximately 70% of patients surviving one year after their fracture, and 7% of patients dying within 30 days of admission (4622/64 858).<sup>1</sup> Numerous patient factors are associated with mortality after hip fracture. It is important that such case-mix variation is considered in the analysis and interpretation of hospital-level results in national clinical audits, not least when this information is used to benchmark services.

The National Hip Fracture Database (NHFD) has been publishing information on patient outcomes following hip fracture in England, Wales and Northern Ireland since 2007. Hospital figures for 30-day mortality rates have always been risk-adjusted by the NHFD, with periodic review of evidence on prognostic models for hip fracture patients to support improvement of its risk adjustment method. This latest refinement to the NHFD risk adjustment model was first implemented in 2014.<sup>2</sup> The NHFD model features the same six patient factors as the model developed by Holt et al<sup>3</sup> from the Scottish Hip Fracture Audit, although different categories are used for some variables.

Prognostic tools specific to patients with hip fracture are unsurprisingly more reliable than generic models,<sup>4,5</sup> but their performances are affected by characteristics of the populations from which they were derived. Different models incorporate a variety of patient characteristics,<sup>3-7</sup> but it is not always clear whether the additional burden of data collection for more complex models is justified by enhanced predictive performance.

The Nottingham Hip Fracture Score is one of the most frequently used and extensively validated outcome prediction models for hip fracture patients.<sup>6-8</sup> It was developed in an English hospital setting and, in 2014, over half (51.7%) of hospitals in England, Wales and Northern Ireland (93/180) reported occasional or routine use of this tool.<sup>9</sup> It has been validated in the United Kingdom and elsewhere against generic models for post-operative mortality such as the Surgical Outcome Risk Tool (SORT) and hip-specific models such as the Almelo Hip Fracture Score (AHFS).<sup>6,8,10-12</sup> It has been recalibrated since its initial development and the creators have recommended

further adjustments to account for changes in the hip fracture population.<sup>8</sup> In this study, the risk model used by the NHFD for 30-day mortality<sup>2,9,13</sup> was evaluated by comparing its performance against the Nottingham model.

## Patients and Methods

**Data sources.** The study used data collected by the NHFD as part of the Royal College of Physicians and Association of Anaesthetists of Great Britain and Ireland collaborative 'sprint' audit on anaesthetic practice. The Anaesthetic Sprint Audit of Practice (ASAP) ran in parallel with standard NHFD data collection over three months in 2013 at 95 hospitals in England, Wales and Northern Ireland.<sup>14</sup> During this time, hospitals submitted data on additional audit fields beyond the standard NHFD dataset, which enabled the estimation of individual patients' risk of 30-day mortality using both NHFD and Nottingham models.

Ethical approval was not required since the NHFD has Section 251 approval from the Health Research Authority's Confidentiality Advisory Group to collect details of hip fracture patients and link their data with the Office for National Statistics (ONS) death register.

**Patient population.** The study included patients who were admitted for hip fracture, between 1 May 2013 and 31 July 2013, to a NHS hospital in England, Wales or Northern Ireland that participated in ASAP. Patients with known operation type were selected for analysis if they were aged between 60 and 110 years, so long as their mortality status at 30 days after surgery could be confirmed by linking their record with the ONS death register.

**The risk models.** The NHFD model contains six variables:<sup>2</sup> age; gender; American Society of Anesthesiologists (ASA) physical status grade; ability to walk indoors; fracture type; and source of admission. The Nottingham model contains seven variables: age; gender; number of comorbidities; abbreviated mental test score on admission; haemoglobin concentration; living in an institution; and malignant disease. The Nottingham model was developed with age stratified into three categories (< 66, 66 to 85, and ≥ 86 years). This study however used the slightly different categories that were applied in the ASAP study.<sup>7</sup> These corresponded to the categories used in the NHFD model. There was also some overlap between the two models in the other patient risk factors that they contain (Table 1).

**Statistical analysis.** We used a standard approach for assessing and improving prediction models<sup>15</sup> whereby a model is refined in one dataset before its performance is examined in a different group of patients.

We used a 'development' dataset of patients admitted between 1 May 2013 and 15 June 2013 to recalculate (recalibrate) the coefficients of the risk factors in the NHFD

**Table I.** Patient factors in the Nottingham and National Hip Fracture Database (NHFD) models

| Patient factor           | In NHFD | NHFD categories   | In Nottingham model | Nottingham categories                         |
|--------------------------|---------|---|---------------------|---|
| Age (yrs)                | Yes     | 60 to 69<br>70 to 79<br>80 to 89<br>90 to 110   | Yes                 | 60 to 69<br>70 to 79<br>80 to 89<br>90 to 110 |
| Gender                   | Yes     | Male<br>Female  | Yes                 | Male<br>Female                                |
| Source of admission      | Yes     | Own home/sheltered housing<br>Not from own home   | N/A                 |   |
| ASA grade                | Yes     | 1 or 2<br>3<br>4 or 5   | N/A                 |   |
| Walking indoors ability  | Yes     | Regularly walked without aids<br>Regularly walked with one aid, two aids or frame<br>Wheelchair or bedbound | N/A                 |   |
| Fracture type            | Yes     | Intracapsular<br>Extracapsular (or other)   | N/A                 |   |
| Number of comorbidities  | N/A     |   | Yes                 | 0 to 1<br>≥ 2                                 |
| AMTS on admission        | N/A     |   | Yes                 | Number from 0 to 10                           |
| Hb on admission          | N/A     |   | Yes                 | Numeric value                                 |
| Living in an institution | N/A     |   | Yes                 | Yes<br>No                                     |
| Malignant disease        | N/A     |   | Yes                 | Yes<br>No                                     |

ASA, American Society of Anesthesiologists; AMTS, Abbreviated Mental Test Score; Hb, Haemoglobin concentration; NHFD, National Hip Fracture Database; N/A, Not Applicable

and Nottingham models. Each patient's probability of death within 30 days of surgery for hip fracture could then be predicted from the regression coefficients of the two models.<sup>16,17</sup>

The resulting risk equations for the two models were applied to the 'validation' dataset of patients admitted between 16 June 2013 and 31 July 2013 to examine their predicted risk of death.<sup>16,17</sup>

The 'discrimination' of a model describes its ability to differentiate between patients who survived or died using the area under the receiver-operating characteristic (ROC) curve. We interpreted values of this 'c-statistic' as indicating poor model performance if they were below 0.70, as moderate for 0.70 to 0.79, and as good performance for 0.80 to 0.89.<sup>12,18</sup>

The 'calibration' of a model compares the predicted and observed mortality rates between groups of patients who are grouped based on their predicted mortality risk. We assessed calibration visually and with the Hosmer-Lemeshow test, with patients allocated to eight predicted risk groups. Goodness of fit was considered adequate if the p-value was less than 0.05.<sup>19</sup>

We addressed missing data using multiple imputation by chained equations (MICE), that is, missing data assumed to be missing at random. Rubin's rules were applied to produce 20 imputed datasets.<sup>20</sup> All analyses were carried out using STATA version 14.1 (StataCorp LP, College Station, Texas).

## Results

During the study period, 8290 patients with hip fracture presented to hospitals that participated in ASAP. The following exclusions were made: n = 116 missing the date of operation; n = 96 missing the type of operation; n = 172 where survival after operation was unknown; and n = 1 with missing data for six risk factors. After exclusions, 7905 patients from 94 hospitals were included in the analysis (Fig. 1).

The distribution of patient factors was relatively similar in the development and validation datasets (Table II). The overall mortality rate was 5.89% (466/7905). The mortality rate within 30 days of surgery for hip fracture was 6.40% in the development dataset (259/4044), but was slightly lower at 5.36% in the validation dataset (207/3861).

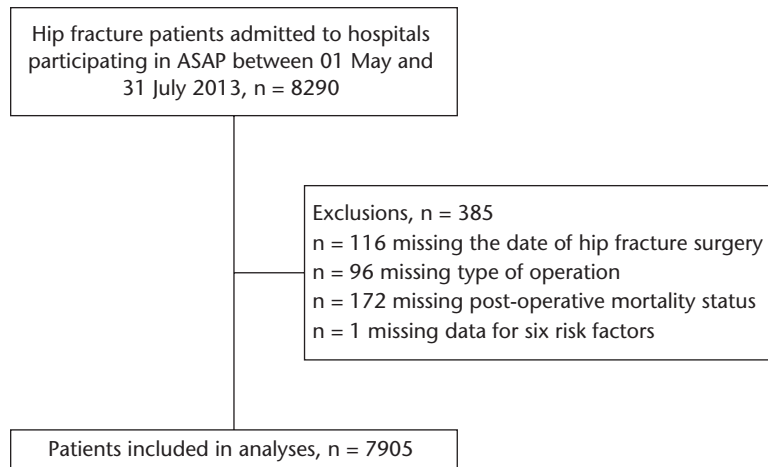


Fig. 1

Flow diagram of study participation. ASAP, Anaesthesia Sprint Audit of Practice.

**Table II.** Patient factors and 30-day mortality rate in development and validation datasets

| Patient factor                                     | Full dataset |      |      | Development set, n = 4044 |      |      | Validation set, n = 3861 |      |      |
|--|--------------|------|------|---------------------------|------|------|--------------------------|------|------|
|  | All patients |      | Died | All patients              |      | Died | All patients             |      | Died |
|  | n            | %    | %    | n                         | %    | %    | n                        | %    | %    |
| <b>Age (yrs)</b>                                   |              |      |      |                           |      |      |                          |      |      |
| 60 to 69   | 711          | 9.0  | 3.1  | 373                       | 9.2  | 2.9  | 338                      | 8.8  | 3.3  |
| 70 to 79   | 1780         | 22.5 | 4.0  | 871                       | 21.5 | 4.8  | 909                      | 23.5 | 3.2  |
| 80 to 89   | 3673         | 46.5 | 5.6  | 1906                      | 47.1 | 5.9  | 1767                     | 45.8 | 5.3  |
| 90 to 110  | 1741         | 22.0 | 9.6  | 894                       | 22.1 | 10.5 | 847                      | 21.9 | 8.6  |
| <b>Gender</b>                                      |              |      |      |                           |      |      |                          |      |      |
| Female   | 5719         | 72.3 | 4.6  | 2940                      | 72.7 | 4.7  | 2779                     | 72.0 | 4.5  |
| Male   | 2186         | 27.7 | 9.3  | 1104                      | 27.3 | 10.9 | 1082                     | 28.0 | 7.7  |
| <b>Source of admission</b>                         |              |      |      |                           |      |      |                          |      |      |
| Own home/sheltered housing                         | 5996         | 75.9 | 4.9  | 3065                      | 75.8 | 5.3  | 2931                     | 75.9 | 4.4  |
| Not from own home                                  | 1909         | 24.1 | 9.1  | 979                       | 24.2 | 9.8  | 930                      | 24.1 | 8.4  |
| <b>ASA grade</b>                                   |              |      |      |                           |      |      |                          |      |      |
| 1 or 2   | 2482         | 31.4 | 1.8  | 1246                      | 30.8 | 1.8  | 1236                     | 32.0 | 1.8  |
| 3  | 4264         | 53.9 | 6.5  | 2197                      | 54.3 | 7.0  | 2067                     | 53.5 | 6.0  |
| 4 or 5   | 954          | 12.1 | 13.8 | 488                       | 12.1 | 15.6 | 466                      | 12.1 | 12.0 |
| Unknown/missing                                    | 205          | 2.6  | 5.4  | 113                       | 2.8  | 6.2  | 89                       | 2.3  | 4.3  |
| <b>Walking indoors ability</b>                     |              |      |      |                           |      |      |                          |      |      |
| Without aids                                       | 3784         | 47.9 | 4.0  | 1926                      | 47.6 | 4    | 1858                     | 48.1 | 4.0  |
| One aid, two aids or frame, wheelchair or bedbound | 4003         | 50.6 | 7.5  | 2055                      | 50.8 | 8.4  | 1948                     | 50.5 | 6.5  |
| Unknown/missing                                    | 118          | 1.5  | 11.9 | 63                        | 1.6  | 12.7 | 55                       | 1.4  | 10.9 |
| <b>Fracture type</b>                               |              |      |      |                           |      |      |                          |      |      |
| Intracapsular                                      | 4563         | 57.7 | 6.1  | 2315                      | 57.2 | 6.5  | 2248                     | 58.2 | 5.7  |
| Extracapsular, including other                     | 3342         | 42.3 | 5.6  | 1729                      | 42.8 | 6.3  | 1613                     | 41.8 | 4.8  |
| <b>Comorbidities</b>                               |              |      |      |                           |      |      |                          |      |      |
| 0 to 1   | 3495         | 44.2 | 3.8  | 1807                      | 44.7 | 3.7  | 1688                     | 43.7 | 4.0  |
| ≥ 2  | 3222         | 40.8 | 7.7  | 1744                      | 43.1 | 8.3  | 1478                     | 38.3 | 6.9  |
| Missing  | 1188         | 15.0 | 7.2  | 493                       | 12.2 | 9.5  | 695                      | 18.0 | 5.5  |
| <b>AMTS on admission</b>                           |              |      |      |                           |      |      |                          |      |      |
| 0 to 6   | 2330         | 29.5 | 8.9  | 1219                      | 30.1 | 9.4  | 1111                     | 28.8 | 8.5  |
| 7 to 10  | 5047         | 63.8 | 4.0  | 2552                      | 63.1 | 4.6  | 2495                     | 64.6 | 3.4  |
| Unknown/missing                                    | 528          | 6.7  | 10.4 | 273                       | 6.8  | 5.3  | 221                      | 5.7  | 11.3 |
| <b>Hb on admission</b>                             |              |      |      |                           |      |      |                          |      |      |
| > 10 g/dl <sup>-1</sup>                            | 7354         | 93   | 5.5  | 3732                      | 92.3 | 5.9  | 3622                     | 93.8 | 5.2  |
| ≤ 10 g/dl <sup>-1</sup>                            | 551          | 7.0  | 10.7 | 312                       | 7.7  | 12.8 | 239                      | 6.2  | 7.9  |
| <b>Living in an institution</b>                    |              |      |      |                           |      |      |                          |      |      |
| No   | 5996         | 75.9 | 4.9  | 3065                      | 75.8 | 5.3  | 2931                     | 75.9 | 4.4  |
| Yes  | 1909         | 24.1 | 9.1  | 979                       | 24.2 | 9.8  | 930                      | 24.1 | 8.4  |
| <b>Malignant disease</b>                           |              |      |      |                           |      |      |                          |      |      |
| No   | 7134         | 90.2 | 5.4  | 3638                      | 90   | 5.9  | 3496                     | 90.5 | 5.0  |
| Yes  | 771          | 9.8  | 10.2 | 406                       | 10   | 11.3 | 365                      | 9.5  | 9.0  |

ASA, American Society of Anesthesiologists; AMTS, Abbreviated mental test score; Hb, Haemoglobin concentration

**Table III.** Performance of Nottingham and NHFD models

| Dataset     | Model      | Patients, n | Hosmer-Lemeshow test |    |         | Area under ROC curve |              |
|-------------|------------|-------------|----------------------|----|---------|----------------------|--------------|
|             |            |             | chi-squared          | df | p-value | C-statistic          | 95% CI       |
| Development | Nottingham | 4044        | 6.27                 | 6  | 0.394   | 0.73                 | 0.70 to 0.76 |
|             | NHFD       | 4044        | 11.3                 | 5  | 0.078   | 0.74                 | 0.71 to 0.77 |
| Validation  | Nottingham | 3861        | 6.10                 | 6  | 0.599   | 0.70                 | 0.67 to 0.74 |
|             | NHFD       | 3861        | 14.9                 | 5  | 0.029   | 0.71                 | 0.68 to 0.75 |

CI, confidence intervals; df, degrees of freedom; NHFD, National Hip Fracture Database; Nottingham, Nottingham Hip Fracture Score; ROC, receiver operating characteristic

**Model recalibration and validation.** The development dataset was used to derive coefficients for the NHFD and Nottingham models and produce the following risk scores:

NHFD risk score =  $\exp(-4.723 + (0.435 \times \text{age } 70 \text{ to } 79) + (0.518 \times \text{age } 80 \text{ to } 89) + (1.060 \times \text{age } 90 \text{ to } 110) + (0.894 \times \text{male}) + (-0.275 \times \text{not admitted from home}) + (1.164 \times \text{ASA } 3) + (1.933 \times \text{ASA } 4 \text{ or } 5) + (0.358 \times \text{walk with at least one aid}) + (-0.055 \times \text{extracapsular fracture}))$ .

Nottingham risk score =  $\exp(-3.955 + (-0.407 \times \text{age } 70 \text{ to } 79) + (0.485 \times \text{age } 80 \text{ to } 89) + (1.083 \times \text{age } 90 \text{ to } 110) + (0.883 \times \text{male}) + (0.558 \times \text{Hb} \leq 10 \text{ g dl}^{-1}) + (0.662 \times \text{cancer}) + (-0.537 \times \text{AMTS} \geq 7) + (0.263 \times \text{living in institution}) + (0.755 \times \geq 2 \text{ comorbidities}))$ .

The NHFD and Nottingham models displayed similar levels of performance (Table III). Both were moderately able to distinguish between patients at high and low risk of death within 30 days, with a c-statistic of 0.71 for the NHFD model and 0.70 for the Nottingham model in the validation dataset.

Model calibration was similar in both models, with a slightly reduced range of predicted risk in the validation dataset (Fig. 2). The Nottingham model displayed a better fit (p-value = 0.599 for Hosmer-Lemeshow test) than the NHFD model (p-value = 0.029), but both were inconsistent in their estimations of mortality risk for patients in the eight risk groups (Fig. 2).

## Discussion

**Main findings.** Case-mix adjustment is important in the development and refinement of prediction models, as potential confounding caused by differences in the populations of individual hospitals must be addressed. The risk model used by the NHFD to adjust 30-day mortality rates for individual hospitals features the same six patient factors as previous NHFD risk models, but with different categories for some variables, and is based on a regression equation rather than a classification tree.<sup>21</sup>

In this evaluation of the NHFD risk model, its performance was comparable with one of the most commonly used and well validated outcome prediction tools for hip fracture – the Nottingham Hip Fracture Score. Both models displayed moderate discriminative power, with the NHFD model achieving a slightly higher c-statistic of 0.71 in the validation dataset. The two models also displayed

moderate levels of calibration, although there was some inconsistent estimation of mortality risk for patients in different risk groups when applied to the validation dataset. These findings are consistent with studies that have found limited goodness of fit for the Nottingham model and the model described by Holt et al.<sup>5,10</sup>

**Strengths and limitations.** Previous studies have excluded certain patient groups, such as those who underwent total hip arthroplasty,<sup>5</sup> but this study included patients who had any type of surgery for hip fracture. The findings could therefore be applied to the whole hip fracture population. The study has several other strengths. It used better quality data than earlier validation studies of the Nottingham model because a high level of data completeness was achieved in the ASAP, where hospitals were excluded if their case ascertainment was 80% or less. We adopted the more robust approach of multiple imputation to manage missing data,<sup>22</sup> while other validation studies limited themselves to complete case analysis by excluding records with missing data.<sup>6,10</sup>

Validation of the Nottingham model using the original and subsequently recalibrated risk score was not feasible as the regression coefficients used to calculate the current Nottingham Hip Fracture Score are not publicly available.<sup>8</sup> Comorbidities were only recorded in the ASAP dataset in terms of whether they were present or not, so it was not possible to determine whether patients without specific recorded comorbidities truly did not have those conditions.

Finally, both models include a selection of patient characteristics associated with death following surgery for hip fracture,<sup>23</sup> but they do not capture all potential factors of interest. Future refinement of either model should be guided by the availability and reliability of recorded variables.

**Implications for further research and clinical practice.** This study has demonstrated that the NHFD model performs comparably with the Nottingham model, and meets current standards for hip fracture outcome prediction in the United Kingdom hip fracture population. Therefore, the NHFD model could be considered for use as an alternative to the Nottingham model, especially if variables (such as haemoglobin concentration or the nature and number of comorbidities) are poorly recorded or missing.



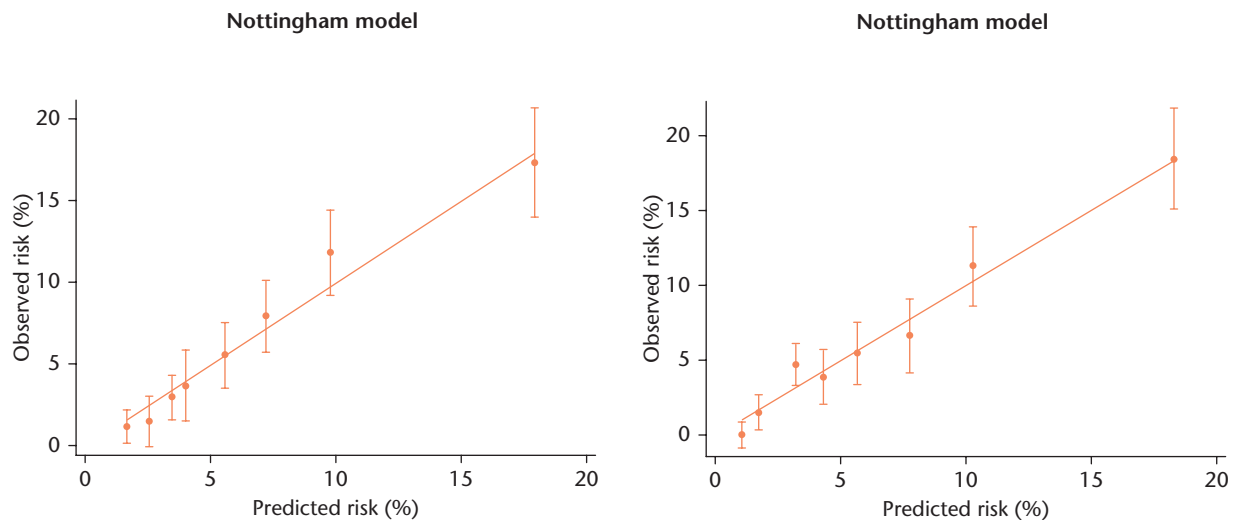


Fig. 2a

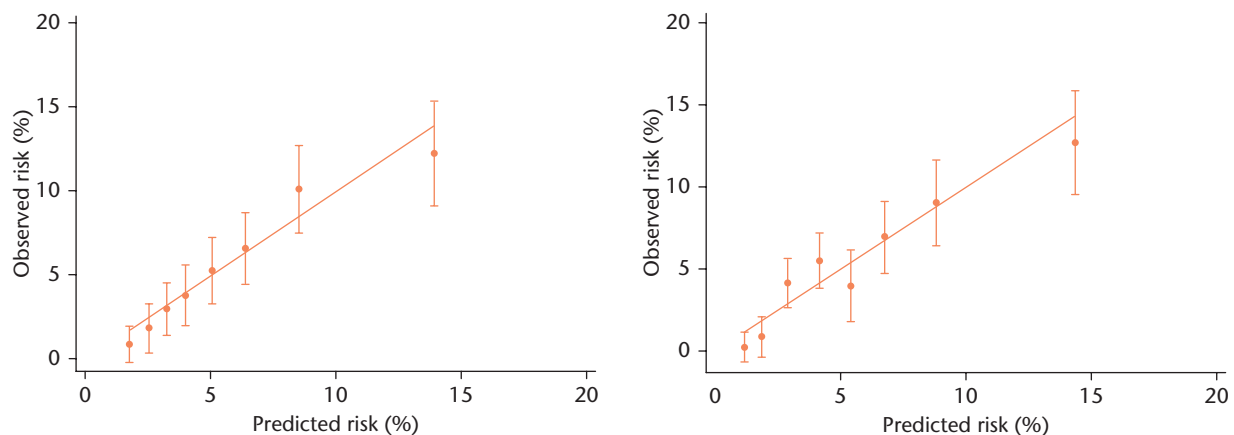


Fig. 2b

Observed *versus* predicted 30-day mortality by risk group, Nottingham and National Hip Fracture Database (NHFD) models. Figure 2a) shows the development dataset, while b) shows the validation dataset.

There is room for improvement in both models. When used to calculate risk-adjusted outcomes of hospitals, there is a small chance of wrongly classifying a hospital as an outlier. However, compared with making predictions for individual patients, the size of the error will be small. This is because the risk adjustment process aims to estimate the average risk of all patients treated at different hospitals, and the differences in the average risks among hospitals is much less than the differences in the risks among individual patients.

If these models are to be used to predict mortality at the individual patient level, we suggest that they be further refined and validated. For example, regular recalibration of both models is appropriate given the trend in decreasing 30-day mortality rate over recent years (7.1% in 2015 compared with 8.0% in 2013).<sup>1,13</sup> Similarly, changes to the profile of hip fracture patients will need to be taken into account in future updates to the models.<sup>24</sup> However, the risk of death faced by an individual patient cannot be fully defined by any model

based on risk factors present on admission, since it will obviously also depend on the quality of care they subsequently receive. The importance of prompt surgery and peri-operative orthogeriatric assessment are just two examples of aspects of care that impact on patients' mortality risk.

While mortality is the most commonly measured patient outcome, there are many other outcomes relevant to the assessment of care delivered to hip fracture patients. For example, functional independence and quality of life in the intermediate future following hip fracture are of interest and importance to patients and their families, care providers and policy makers. Future studies would benefit from assessing the performances of these tools in their ability to predict differences among patients on these outcomes. This would lead to a better understanding of how different patient groups (such as those at relatively high risk or low risk of death within 30 days of surgery) are affected by hip fracture in the months after completing hospital treatment.

## References

- No authors listed.** National Hip Fracture Database, Annual Report 2016. <https://www.rcplondon.ac.uk/projects/outputs/national-hip-fracture-database-annual-report-2016> (date last accessed 28 June 2017).
- Tsang C, Cromwell D.** Statistical methods developed for the National Hip Fracture Database annual report, 2014: a technical report. <https://www.nhfd.co.uk> (date last accessed 23 May 2017).
- Holt G, Smith R, Duncan K, Finlayson DF, Gregori A.** Early mortality after surgical fixation of hip fractures in the elderly: an analysis of data from the Scottish hip fracture audit. *J Bone Joint Surg [Br]* 2008;90-B:1357-1363.
- Marufu TC, Mannings A, Moppett IK.** Risk scoring models for predicting peri-operative morbidity and mortality in people with fragility hip fractures: qualitative systematic review. *Injury* 2015;46:2325-2334.
- Karres J, Heesakkers NA, Ultee JM, Vrouwenraets BC.** Predicting 30-day mortality following hip fracture surgery: evaluation of six risk prediction models. *Injury* 2015;46:371-377.
- Wiles MD, Moran CG, Sahota O, Moppett IK.** Nottingham Hip Fracture Score as a predictor of one year mortality in patients undergoing surgical repair of fractured neck of femur. *Br J Anaesth* 2011;106:501-504.
- Maxwell MJ, Moran CG, Moppett IK.** Development and validation of a preoperative scoring system to predict 30 day mortality in patients undergoing hip fracture surgery. *Br J Anaesth* 2008;101:511-517.
- Moppett IK, Parker M, Griffiths R, et al.** Nottingham Hip Fracture Score: longitudinal and multi-centre assessment. *Br J Anaesth* 2012;109:546-550.
- No authors listed.** National Hip Fracture Database, Annual Report 2015. <http://www.nhfd.co.uk/nhfd/nhfd2015reportPR1.pdf> (date last accessed 28 June 2017).
- Rushton PRP, Reed MR, Pratt RK.** Independent validation of the Nottingham Hip Fracture Score and identification of regional variation in patient risk within England. *Bone Joint J* 2015;97-B:100-103.
- Nijmeijer WS, Folbert EC, Vermeer M, Slaets JP, Hegeman JH.** Prediction of early mortality following hip fracture surgery in frail elderly: The Almelo Hip Fracture Score (AHFS). *Injury* 2016;47:2138-2143.
- Marufu TC, White SM, Griffiths R, Moonesinghe SR, Moppett IK.** Prediction of 30-day mortality after hip fracture surgery by the Nottingham Hip Fracture Score and the Surgical Outcome Risk Tool. *Anaesthesia* 2016;71:515-521.
- No authors listed.** National Hip Fracture Database, Annual Report 2014. <http://www.nhfd.co.uk/nhfd/nhfd2015reportPR1.pdf> (date last accessed 28 June 2017).
- No authors listed.** Anaesthesia Sprint Audit of Practice 2014. <https://www.aagbi.org/> (date last accessed 28 June 2017).
- Altman DG, Royston P.** What do we mean by validating a prognostic model? *Stat Med* 2000;19:453-473.
- Royston P, Moons KG, Altman DG, Vergouwe Y.** Prognosis and prognostic research: Developing a prognostic model. *BMJ* 2009;338:b604.
- Grant SW, Hickey GL, Grayson AD, Mitchell DC, McCollum CN.** National risk prediction model for elective abdominal aortic aneurysm repair. *Br J Surg* 2013;100:645-653.
- Swets JA.** Measuring the accuracy of diagnostic systems. *Science* 1988;240:1285-1293.
- Lemeshow S, Hosmer DW Jr.** A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 1982;115:92-106.
- White IR, Royston P, Wood AM.** Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;30:377-399.
- No authors listed.** National Hip Fracture Database, National Report 2013. [http://www.nhfd.co.uk/20/hipfractureR.nsf/0/CA920122A244F2ED802579C900553993/\\$file/NHFD%20Report%202013.pdf](http://www.nhfd.co.uk/20/hipfractureR.nsf/0/CA920122A244F2ED802579C900553993/$file/NHFD%20Report%202013.pdf) (date last accessed 28 June 2017).
- Marshall A, Altman DG, Royston P, Holder RL.** Comparison of techniques for handling missing covariate data within prognostic modelling studies: a simulation study. *BMC Med Res Methodol* 2010;10:7.
- Smith T, Pelpola K, Ball M, Ong A, Myint PK.** Pre-operative indicators for mortality following hip fracture surgery: a systematic review and meta-analysis. *Age Ageing* 2014;43:464-471.
- Neuburger J, Wakeman R.** Is the incidence of hip fracture increasing among older men in England? *J Epidemiol Community Health* 2016;70:1049-1050.

### Acknowledgements

- This publication is based on data collected by or on behalf of the Healthcare Quality Improvement Partnership, who have no responsibility or liability for the accuracy, currency, reliability and/or correctness of this publication. We thank the Falls and Fragility Fracture Audit Programme audit team and Crown Informatics for their help to access NHFD data.
- We are grateful to the steering group of the ASAP project, with particular thanks due to Dr Richard Griffiths, Dr Stuart White and Dr Amer Majeed of the Association of Anaesthetists of Great Britain and Ireland for their support throughout and subsequent to the ASAP project.
- We are grateful to Dr Iain Moppett for his guidance on mapping the patient factor of 'living in an institution' to NHFD data.

### Funding Statement

- Carmen Tsang reports that funding has been received from Healthcare Quality Improvement Partnership (HQIP) which is related to this paper.
- The NHFD is commissioned by the Healthcare Quality Improvement Partnership (HQIP), and managed by the Clinical Effectiveness and Evaluation Unit of the Royal College of Physicians (RCP) as part of the Falls and Fragility Fracture Audit Programme (FFFAP).
- The Anaesthesia Sprint Audit of Practice (ASAP) was commissioned by the Healthcare Quality Improvement Partnership and managed by the Royal College of Physicians in partnership with the Association of Anaesthetists of Great Britain and Ireland (AAGBI).

### Author Contribution

- C. Tsang: Study concept and design, Data analysis, Interpretation of the data, Drafting, revising and finalising the paper
- C. Boulton: Study concept and design, Interpretation of the data, Drafting, revising and finalising the paper
- V. Burgon: Study concept and design, Interpretation of the data, Drafting, revising and finalising the paper
- A. Johansen: Study concept and design, Interpretation of the data, Drafting, revising and finalising the paper
- R. Wakeman: Study concept and design, Interpretation of the data, Drafting, revising and finalising the paper
- D. A. Cromwell: Study concept and design, Data analysis, Interpretation of the data, Drafting, revising and finalising the paper

### Conflicts of Interest Statement

- D. Cromwell and C. Tsang were commissioned by the Royal College of Physicians to conduct the statistical analyses for the Falls and Fragility Fracture Audit Programme, which includes the NHFD.
- C. Boulton is Programme Manager of the Falls and Fragility Fracture Audit Programme.
- V. Burgon was Project Manager of the National Hip Fracture Database.
- A. Johansen is Clinical Lead for Orthogeriatrics of the National Hip Fracture Database.
- R. Wakeman is Clinical Lead for Orthopaedic Surgery of the National Hip Fracture Database.

© 2017 Tsang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution licence (CC-BY-NC), which permits unrestricted use, distribution, and reproduction in any medium, but not for commercial gain, provided the original author and source are credited.